PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 7 : H04Q 11/04

A1

(11) International Publication Number: WO 00/30401

(43) International Publication Date: 25 May 2000 (25.05.00)

(21) International Application Number: PCT/US99/26902

(22) International Filing Date: 12 November 1999 (12.11.99)

(30) Priority Data:

60/108,331	13 November 1998 (13.11.98)	US
09/245,792	5 February 1999 (05.02.99)	US

(71) Applicant (*for all designated States except US*): NORTEL NETWORKS CORPORATION [CA/CA]; World Trade Center of Montreal, 8th floor, 380 St. Antoine Street West, Montreal, Quebec H2Y 3Y4 (CA).

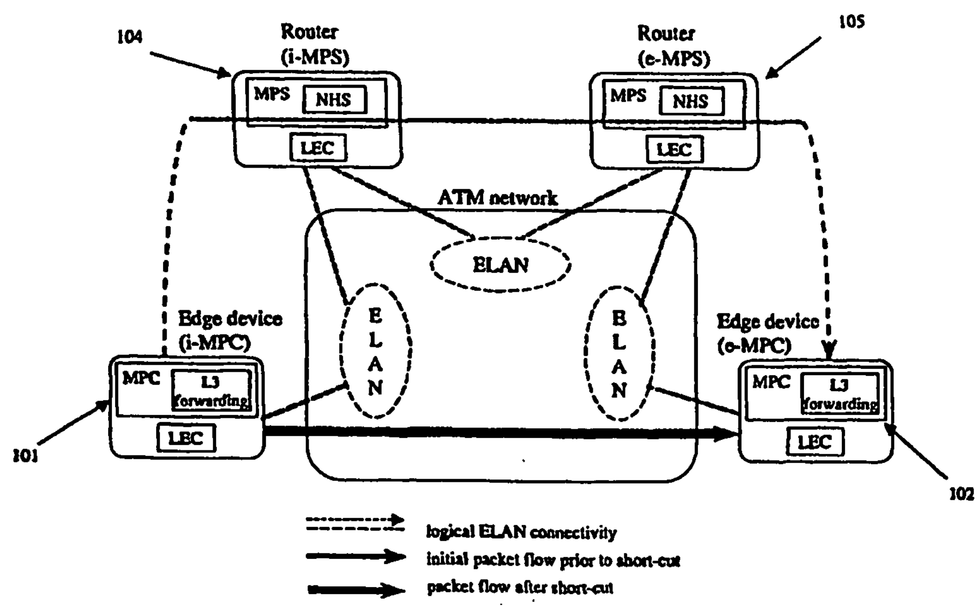
(72) Inventors; and
(75) Inventors/Applicants (*for US only*): HSU, Ivy [US/US]; P.O. Box 58185, Santa Clara, CA 90025-1026 (US). SQUIRE, Matthew [US/US]; P.O. Box 58185, Santa Clara, CA 90025-1026 (US). BOTTORFF, Paul [US/US]; P.O. Box 58185, Santa Clara, CA 90025-1026 (US).

(74) Agents: SCHAAL, William, W. et al.; Blakely, Sokoloff, Taylor & Zafman, 7th Floor, 12400 Wilshire Boulevard, Los Angeles, CA 90025-1026 (US).

(81) Designated States: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published
With international search report.
Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(54) Title: METHOD AND APPARATUS FOR SUPPORT OF IP DIFFERENTIATED SERVICE OVER MPOA



(57) Abstract

A method and apparatus for providing support of IP differentiated service over MPOA. In the described embodiment, methods are described for distribution of policy information to MPOA clients, for flow detection by MPOA clients and for discovery of differentiated service capability of MPOA clients.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania
AM	Armenia
AT	Austria
AU	Australia
AZ	Azerbaijan
BA	Bosnia and Herzegovina
BB	Barbados
BE	Belgium
BF	Burkina Faso
BG	Bulgaria
BJ	Benin
BR	Brazil
BY	Belarus
CA	Canada
CF	Central African Republic
CG	Congo
CH	Switzerland
CI	Côte d'Ivoire
CM	Cameroon
CN	China
CU	Cuba
CZ	Czech Republic
DE	Germany
DK	Denmark
EE	Estonia
ES	Spain
FI	Finland
FR	France
GA	Gabon
GB	United Kingdom
GE	Georgia
GH	Ghana
GN	Guinea
GR	Greece
HU	Hungary
IE	Ireland
IL	Israel
IS	Iceland
IT	Italy
JP	Japan
KE	Kenya
KG	Kyrgyzstan
KP	Democratic People's Republic of Korea
KR	Republic of Korea
KZ	Kazakstan
LC	Saint Lucia
LI	Liechtenstein
LK	Sri Lanka
LR	Liberia
LS	Lesotho
LT	Lithuania
LU	Luxembourg
LV	Latvia
MC	Monaco
MD	Republic of Moldova
MG	Madagascar
MK	The former Yugoslav Republic of Macedonia
ML	Mali
MN	Mongolia
MR	Mauritania
MW	Malawi
MX	Mexico
NE	Niger
NL	Netherlands
NO	Norway
NZ	New Zealand
PL	Poland
PT	Portugal
RO	Romania
RU	Russian Federation
SD	Sudan
SE	Sweden
SG	Singapore
SI	Slovenia
SK	Slovakia
SN	Senegal
SZ	Swaziland
TD	Chad
TG	Togo
TJ	Tajikistan
TM	Turkmenistan
TR	Turkey
TT	Trinidad and Tobago
UA	Ukraine
UG	Uganda
US	United States of America
UZ	Uzbekistan
VN	Viet Nam
YU	Yugoslavia
ZW	Zimbabwe

METHOD AND APPARATUS FOR SUPPORT OF IP DIFFERENTIATED SERVICE OVER MPOA

CROSS REFERENCE TO RELATED APPLICATIONS

This application claims benefit of co-pending U.S. provisional application Serial No. ______________ titled "Method and Apparatus for Support of IP Differentiated Services over MPOA" filed November 13, 1998.

BACKGROUND OF THE INVENTION

Developing Quality of Service (QoS) capabilities for Internet Protocol (IP) has been an active work area at the Internet Engineering Task Force (IETF) as well as by many other entities. The IETF has identified two alternative approaches, Integrated Services (Int-Serv) and Differentiated Services (Diff-Serv) as the most promising solutions to address different QoS needs.

The IETF has been addressing the service mappings between Int-Serv and individual link layers, including ATM (see, e.g., L. Berger, "RSVP over ATM Implementation Guidelines", IETF RFC 2379, August 1998 and M. Garrett and M. Borden, "Interoperation of Controlled-Load Service and Guaranteed Service with ATM", IETF RFC 2381, August 1998.) More recently, a proposal was accepted by the ATM Forum to commence work in the Traffic Management working group on enhancements for supporting both Diff-Serv and IEEE 802.1D (see, e.g., Marty Borden et al., "Enhancements to Support IETF Diff-Serv and IEEE802.1D", ATM Forum, atmf/98-0789R1, Oct. 1998.)

Unfortunately, neither of these proposals are intended to address support of Diff-Serv using MPOA. Diff-Serv flows over standard MPOA do not receive proper service because, as will be explained in greater detail below, shortcut flows to a given destination are forwarded over the same ATM connection.

These shortcuts bypass intermediate routers that would otherwise be responsible for providing different services to different flows. Therefore, it is necessary to distribute policy information to the various MPOA clients. The policy information defines how packets should be classified into Diff-Serv classes.

These proposals further do not adequately address flow detection by the MPOA clients

Further, the proposals do not adequately address discovery of Diff-Serv capabilities of MPOA devices.

What is needed in order to effectively implement Diff-Serv in a joint IP/ATM environment is a method and apparatus for providing distribution of policy information to MPOA clients, method and apparatus for addressing flow detection requirements of MPOA clients, and method and apparatus providing discovery of Diff-Serv capabilities of MPOA devices.

Before addressing solutions proposed by the present invention, it is worthwhile to provide some background on MPOA and Diff-Serv.

MPOA

MPOA establishes direct shortcuts across an ATM network for forwarding Internetwork Layer packets. Shortcuts take advantage of the ATM topology and can be more efficient than hop-by-hop router forwarding. This is illustrated by Figure 1 which shows logical components and packet flows for an MPOA network. As can be seen, in a network not implementing MPOA, packet flow follows the logical ELAN connectivity route in order to transport a packet from edge device 101 to edge device 102. The actual route is shown (without MPOA

shortcuts) is shown as passing through routers 104 and 105. The route, after the shortcut is established using MPOA, is represented by the line passing through the ATM network.

MPOA requires six distinct operations: (1) configuration, (2) discovery, (3) flow detection, (4) target resolution, (5) connection management and (6) data transfer. MPOA devices obtain configuration via a LAN Emulation Configuration Server (LECS). Each of these operations will be discussed in greater detail below.

MPOA configuration information, as well as LANE configuration information, is returned in the LAN Emulation (LANE) configuration process. The MPOA configuration information is needed to initialize and control other MPOA operations. MPOA configuration information consists of the internetwork protocols monitored by the MPOA device, timeouts, and thresholds.

Discovery is the process by which MPOA components learn of each other's existence. Additional information is included in LANE control frames to permit automatic detection of MPOA components, thus eliminating some configuration and making the environment more dynamic. Discovered information consists of the type of device (either an MPOA Client which will be referred to as a "MPC" or a MPOA server referred to as a "MPS") and the addresses used by the device.

Flow detection is performed by MPCs to identify streams of traffic, or flows, that should be transmitted over a shortcut. The default definition of a flow in MPOA is a sequence of packets to a particular internetwork destination that

satisfies a certain transmission rate. Other (unspecifi d) flow definitions are permitted.

After a flow is detected, target resolution is the process by which an MPC determines the mapping between an internetwork address and an ATM address. A resolution request is forwarded along the routed path from the Ingress MPC (I-MPC) to the egress MPC (E-MPC), which returns a response indicating the ATM address to be used for a shortcut for that destination. Additional information, including the encapsulation for the shortcut packets, is also returned in the response.

Connection management then establishes a Virtual Circuit Connection (VCC) to this address for the purposes of transmitting packets to the destination. The default QoS for MPOA shortcuts is best effort Unspecified Bit Rate (UBR) (although other types of QoS are permitted).

Data transfer refers to the transmission of packets over the shortcut. These packets bypass all intermediate routers, taking a more direct path (the shortcut) over the ATM network. MPOA v1.0 was developed to support a best-effort service at the internetwork layer. Thus, it lacked the specific features necessary to support QoS at the internetwork layer, such as those characterized by Diff-Serv and Int-Serv.

Differentiated Service

Diff-Serv differs from Int-Serv in a number of ways which have implications for MPOA and which have been heretofore unaddressed. In particular:

- Service categories in th Diff-Serv model are relative, or qualitative, in that the aggregate of packet flows with the same Diff-Serv Code Point (DSCP) is subject to the same Per-Hop Behavior (PHB), which may not involve explicit resource commitment per flow.

- In Diff-Serv, the end-systems need not explicitly signal the packet classification and service requirement to the network. This is unlike the use of PATH and RESV messages in Resource ReSerVation Protocol ("RSVP") which is used to support Int-Serv. Thus, RSVP over MPOA will require treatment of these signaling messages, while Diff-Serv does not.

- Diff-Serv uses an implicit, policy-based model. Diff-Serv services are constructed by a combination of per hop behaviors (PHBs) and edge behaviors. A Diff-Serv domain consists of *interior* and *edge* nodes, with most of the complexity at the edge nodes. Interior nodes examine the DS field only (known as Behavior Aggregate classifier) and perform the appropriate PHB. In addition to PHBs, edge nodes may be responsible for classification, metering, policing, marking, and shaping (among other functions). The edge actions are determined based on the value of a combination of one or more header fields. The rules for each service are set through administrative policy. Policy can be distributed to Diff-Serv nodes via a policy protocol.

The IETF Differentiated Services (DS) Working Group is developing a way of providing Intemet Protocol (IP) QoS. DS uses the TOS octet in IPv4 and the

Traffic Class octet in IPv6, together referred to as the DS field, to indicate the QoS class that an IP packet belongs to. This enables service discrimination without requiring per-flow states and signaling at routers.

A DS code point is a specific pattern of the DS field. Associated with a DS code point is a per-hop behavior (PHB), which is the packet forwarding treatment that a DS-compliant node should apply at its output interface to packets marked with this DS code point.

A DS domain consists of *interior* and *edge* nodes, with most of the complexity at the edge nodes. Interior nodes examine the DS field and perform the appropriate PHB. In addition to PHBs, edge nodes may be responsible for classification, metering, policing, marking, and shaping (among other functions). A DS edge device performs *classification* by inspecting the header of each packet and considering it part of a specific group or class. An edge device uses *metering* to measure the data rate in a particular class. *Policing* is the means by which a specific data rate is enforced. *Marking* is when a DS edge device fills in the DS field of a packet based on the class to which the packet belongs. Finally, an edge device performs *shaping* by transmitting packets at specific rates for various traffic classes. A particular edge device may be required to do any subset of these functions (including none).

DS services are constructed by a combination of per-hop behaviors and edge behaviors. The rules for each service are set through administrative policy. DS policy has several aspects. All nodes must know the relationship between the DS field of a packet and the appropriate PHB. When transmitting over links with QoS abilities such as ATM, this may include a mapping between the IP DS code point and specific link layer QoS parameters. Additionally, edge nodes must know the rules for packet classification, metering, policing, marking, and shaping.

Policy is distributed to DS nodes via a policy protocol. The exact form of the protocol is as yet undefined.

One prevalent set of possible DS services is soft or relative QoS for aggregate flows, which is referred to here as a Class of Service (CoS) differentiation. Contrary to the Resource Reservation Protocol (RSVP) approach, flow identifications are not signaled explicitly by the end systems but are defined by policy. For example, an output interface may contain a set of parallel queues that are served based on a Weighted Round Robin (WRR) scheme. Each DS code point is then associated with one queue with a serving weight. Edge devices mark packets according to configured policy, so that all packets with the same DS code point share the same forwarding queue and are entitled to its bandwidth share at each output interface. This is just one possible DS interpretation.

Unfortunately, if DS edge and per-hop behaviors are only implemented in routers, then packets forwarded over MPOA shortcuts bypass the routers and may not receive the proper service.

BRIEF SUMMARY OF THE INVENTION

A method and apparatus providing for extensions to MPOA to accommodate Diff-Serv over MPOA is described including: (a) distribution of policy, (b) flow detection and (c) discovery of diff-serv capabilities of MPOA clients.

These and other aspects of the present invention will be better described with reference to the detailed description.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 is an overall diagram of a network as may implement an embodiment of the present invention.

DETAILED DESCRIPTION OF EMBODIMENT(S) OF THE INVENTION

As was discussed in the background section, implementing internet protocol (IP) differentiated services (Diff-Serv) using Multiprotocol Over ATM (MPOA) has heretofore not been addressed.

The present invention introduces extensions to MPOA in the following areas in order to provide for Diff-Serv over MPOA: (a) distribution of policy, (b) flow detection and (c) address discovery.

Policy Distribution

Differentiated Services use policy to determine to control packet flows. In the described embodiment, it is assumed that policy information is distributed to all routers (e.g., router 104 and 105 in figure 1) in the differentiated service domain. This implies that policy is distributed to all MPOA servers (MPSs). However, as has been stated, within the MPOA service domain, the responsibility for forwarding packets is distributed between MPSs 104, 105 and MPOA clients (MPCs) 101, 102. Thus, in order to provide per hop behavior (PHB) service, the MPCs as well as the MPSs must have access to the DS policy.

For interior Diff-Serv nodes, the policy consists of determining the resource allocation and mapping the DS field of a packet into specific ATM QoS parameters. Edge nodes have additional policy considerations. DS policy for edge nodes can be expressed as a collection of filter specifications. A filter specification is a combination of filter criteria and filter actions. Filter criteria define how packets should be classified. They are generally based on certain header fields (e.g., source IP address, destination IP address, carried protocol,

source port, and destination port). Filter actions define how packets should b marked and treated by DS edges.

What is proposed here is to allow the DS routers, in their roles as MPSs, to distribute policy to MPCs. When a router is acting as an interior DS device, it must distribute the resource allocation and the mappings between DS code points and ATM QoS parameters to its MPCs. The policy may be distributed by any of a number of methods including, for example:

1. The MPC can run the policy distribution protocol to receive policy, similar to an MPS.
2. The mappings can be distributed during configuration via the LECS.
3. The mappings can be distributed during discovery.
4. A new request/response/trigger frame between MPCs and MPSs can be defined to distribute this information.

The first method greatly increases the number of policy clients, and may affect the scalability of the policy distribution protocol. However, this method advantageously requires no extensions in the behavior of any other LANE or MPOA component.

In methods (2) through (4), new type length values (TLVs) fields are defined to carry the mappings. The TLVs are carried in the appropriate LANE or MPOA control frames. The second method requires some out-of-band mechanism to distribute policy to the LANE Configuration Server, which then distributes it to the MPCs. The third method is dynamic, but may be limited by the size of the LANE control frame (which at most 1516 octets). The final method is the most generic, but requires a new type of MPOA control frame, a policy request/response frame. With the last method, the MPC must request the mappings after

discovering an MPS and before establishing any shortcuts. Methods (3) and (4) have the advantage that the distribution is dynamic (i.e., the mappings can change and the changes are automatically propagated to the MPCs).

When an MPS is acting as an edge router, it must also distribute filter specifications to MPCs. Filter specifications are required so that the MPC can properly function as a DS edge device. Since shortcut traffic bypasses the MPS, which is also the DS edge router, the MPC must perform the classification, marking, metering, policing, and shaping functions on behalf of the MPS for packets traversing the shortcut. This proposal defines the 4 methods above to distribute filter specifications to MPCs.

The distribution of filter specifications differs from the distribution of the mappings between DS code points and ATM QoS parameters in several aspects. First, the sheer amount of information in the collection of filter specifications is much greater. There are a limited number of DS code points, but the number of possible filter specifications is much greater. Second, it is unlikely that a single MPC will need all of the filter specifications at any time. An MPC needs only those filter specifications that apply to traffic being forwarded by that MPC. Third, filter specifications are generally more dynamic and changes in filter specifications must be propagated in a timely fashion.

Filter specifications can be classified into two categories: destination-specific and destination-independent. Destination independent filters apply to more than one destination IP address. It is important that an MPC have all of filter specifications that apply to a particular destination before the MPC forwards traffic for that destination over a shortcut, so that the appropriate DS behavior can be applied to the shortcut traffic.

Methods (1) through (3) can be used to distribute some or all filter specifications to th_ MPC. However, method (4) permits filter specifications to be distributed to MPCs on an as-needed basis as long as the following requirements are met:

- The MPC must request all filters that apply to a particular destination before deciding to establish a shortcut to the destination.
- The MPS responds to filter requests by returning the applicable filter specifications. The set of applicable filter specifications could be empty.
- The MPS may generate a filter trigger to an MPC. A filter trigger causes the MPC to initiate a filter request for the filter specifications indicated by the trigger. The trigger may be used to indicate a change or update in policy.

These requirements can be satisfied in a variety of ways:

- In some implementations, this may result in an MPC requesting all destination independent filters shortly after discovering an MPS.
- In some implementations, this may result in an MPC requesting all filters that apply to a particular destination after initiating flow detection but before initiating a resolution request. The MPC may signal that it wants all filters that apply to a target (including destination independent filters), or only that it wants those filters that apply to the specific destination.

Other implementations satisfying the requirements are possible. The key point is that the MPC must be guaranteed to have all applicable filter specifications before it attempts to establish a shortcut. The applicable specifications are requested by the MPC, and the MPC may be prodded to request certain filter specifications by the MPS.

MPOA Configuration and Discovery

MPOA components discover each other using extensions to the LANE LE_ARP protocol that carry information such as the MPOA device type and ATM address. In the described embodiment, new TLV(s) are added to the LE_ARP messages to indicate the DS capabilities of the device. The absence of the TLV(s) indicates that the component does not support any of the DS extensions. An MPC or MPS capable of the DS extensions will not attempt to use them with an MPS or MPC not capable of the extensions. This provides for interoperability with current MPOA implementations.

In the described embodiment, TLVs can be used to indicate support of the following DS capabilities: (a) DS parameter distribution, (b) filter specification distribution, (c) whether policy distribution is enabled and (d) whether an MPS is a DS edge router as well as other capabilities. In certain embodiments, indicating other DS abilities via discovery TLVs is also possible.

Flow Detection

In MPOA v1.0, the detection of shortcut-eligible IP flows is, by default, based on the number of packets sent to a particular destination through a particular MPS in a specified period of time. In the described embodiment, the default flow detection is extended to be the number of packets with a particular DS code point sent to a particular destination through a particular MPS in a specified period of time. Other algorithms for flow detection may be utilized in certain embodiments but the flow detection algorithm should use the DS field in the definition of a flow. In particular, ingress cache entries in MPCs must be extended to monitor the DS field.

It is worthwhile to describe two embodiments of the d fault flow detection algorithm. In the first option, DS-sensitive flow detection (if using method (1) or (2) of the previous section) can be initiated upon the creation of the ingress cache entry (i.e. on the detection of the first packet with a particular destination address, MPS, and DS field). In the second option, in which the I-MPC does not have the filters associated with a given destination and hence must obtain the filters dynamically, the I-MPC performs a two-stage flow detection. In the first stage, the I-MPC counts the number of packets with a particular destination address and MPS. When this count exceeds a threshold, the MPC initiates a filter request. Upon receiving the applicable filter specifications, the I-MPC can perform flow detection including the DS field, based on the filter specifications returned.

Other default flow detection algorithms may be utilized without departure from the present invention.

Assume that an MPC is performing the edge device functions for a particular MPS and that it has requested and received the filter specifications applicable to a certain packet. When another matching packet arrives at the I-MPC, it is matched against the filter specifications to determine its DS code point. The result, together with the destination address and MPS, is used to match with the ingress cache entries, or to create a new entry if one does not already exist. If a match is found and a shortcut VCC of the appropriate QoS is available, the I-MPC applies the corresponding filter actions before forwarding the packet over the shortcut. Otherwise, a flow detection counter is incremented. If the configured flow threshold is exceeded, a shortcut with the appropriate QoS class is initiated. If the MPC is not performing edge router functions, the DS code point of the packet is used to match the ingress cache entry.

Target Resolution

Through target resolution, an I-MPC obtains the ATM address of the E-MPC for the destination IP address. MPOA provides for obtaining this information. However, additional information, such as encapsulation and tagging, may be conveyed through target resolution. For example, the E-MPC needs to know the DS code point of the flows that prompt the resolution process in the following cases:

(1) When different DS codepoints need to be mapped to different egress 802.1D user_priority markings in the DLL header. In this case, the layer 2 encapsulation at the egress is dependent on the DS codepoint.

(2) When E-MPC uses different MPOA tags for different DS codepoints to aid egress queue selection at its output interfaces.

Thus, as one aspect of the described embodiment, it is proposed to provide enhancements that may be needed in such environments. An I-MPC can include an extension in the target resolution to indicate the DS codepoint for the flow instigating the request.

Connection Management

Shortcut VCCs for flows with different DS code points may require different ATM QoS capabilities than UBR. Thus, in the described embodiment, associated with each DS code point is a set of ATM signaling Information Elements (IEs), which specify the QoS requirements and traffic parameters that are appropriate for that code point. The logic for determining whether to share an existing VCC or to establish a new one is similar to MPOA v1.0. With DS, the

ATM signaling IEs associated with the flow's DS Code p int should be used for establishing a new VCC.

Further details regarding a proposed embodiment of the present invention are given in Table A below which discusses several issues with implementation of DS over MPOA and proposed solutions:

TABLE A

ISSUES

Issue 1:

With MPOA, the packet forwarding part of a router is distributed to the MPCs for packets that are forwarded over shortcuts. This implies that a router must also share its Diff-Serv roles (per-hop behaviors, edge behaviors) with its MPCs. In other words, MPCs must have access to the relevant Diff-Serv policy. Otherwise shortcut forwarding would continue to use the default UBR VCCs for all packets, bypassing the appropriate Diff-Serv treatments.

Diff-Serv policy has several aspects. All nodes, interior and edge, must know the relationship between the DS field of a packet and the appropriate PHB. When transmitting over links with QoS abilities such as ATM, this may include a mapping between the DS Codepoint (DSCP) and specific link layer QoS parameters. Additionally, edge nodes must know the rules for packet classification, metering, policing, marking, and shaping (the policy decides if an edge node should perform some, all, or none of these functions).

An MPS at a Diff-Serv interior node needs to forward to its MPCs the mapping between the DSCPs and ATM signaling parameters (i.e., QoS and traffic descriptor IEs). Edge policy, on the other hand, can be far more complex. Diff-Serv policy for edge nodes can be expressed as a collection of *filter specifications*. A filter specification is a combination of *filter criteria* and *filter actions*. Filter criteria define how packets should be classified. They are generally based on certain header fields (e.g., source IP address, destination IP address, carried protocol, source port, and destination port). Filter actions define how packets should be marked and treated by edge nodes.

Furthermore, there are two related questions that also need to b addressed: the issue of policy distribution timing (i.e., when a policy must become available at the MPCs) and the issue of policy changes and updates.

Issue 2:

In MPOA v1.0, the default detection of shortcut-eligible IP flows is based on the count of packets sent to a particular destination through a particular MPS in a specified period of time. With Diff-Serv QoS, packets for the same destination may be of different Diff-Serv categories and therefore would require multiple ATM VCCs (e.g., VCCs that are signaled with different Virtual Class Selector IE.) The definitions of flows and shortcut-eligibility must therefore be clarified and enhanced. For example, should packets to the same destination but with different DSCPs be considered as multiple separate flows or one flow? What definitions are chosen would have some implication on the subsequent steps of target resolution and VCC establishment. Although MPOA v1.0 permits other flow detection procedures, interoperability concerns dictate a standard default flow detection method for Diff-Serv capable MPCs.

Issue 3:

With Diff-Serv, an I-MPC may need to convey to the E-MPC the DSCP of the flow (or flows) that prompted the target resolution process. Some of the important applications include:

When different DSCPs need to be mapped to different IEEE 802.1D user_priority markings in the DLL header for frames departing from the E-MPC. In this case, the layer 2 encapsulation at the egress is dependent on the DSCP.

- When an E-MPC wants to use different MPOA tags for different DSCPs to aid egress queue selection at its output interfaces. With different MPOA tags, the E-MPC avoids needing to reexamine the DS field in the packet header for output queue selection.

This requirement suggests that the target resolution process needs to accommodate extensions for Diff-Serv in the control messages. Note that other fields in a packet besides the DSCP may be needed by the egress to determine the correct DLL/tag. Such fields need to be transmitted to the egress during the resolution process.

<u>Issue 4:</u>

Two Diff-Serv domains can be interconnected by a boundary node, which performs any necessary re-marking and/or traffic shaping of packets. When multiple Diff-Serv domains are supported over a single ATM network cloud (e.g., in an enterprise network with logical departmental boundaries to observe different Diff-Serv policies), there is a problem with MPOA shortcuts spanning Diff-Serv domains: packets over the shortcuts bypass the boundary node and do not get reclassified. This requires MPOA to be aware of the Diff-Serv domain boundary, or for the MPOA Diff-Serv solution to be defined for intra-domain differentiated services only.

<u>Issue 5:</u>

In order to continue to support plug-n-play operation, the discovery protocol must be enhanced to support the automatic discovery of the QoS capabilities of neighboring MPOA devices.

PROPOSED SOLUTIONS

<u>Policy Distribution</u>

There are several approaches for an MPC to obtain its relevant DS policy:

1. The MPC can obtain policy in the same manner as its MPS, such as running a policy distribution protocol to receive policy directly from a server, or being configured by the network management system.
2. The policy can be distributed during configuration via the LECS.
3. The policy can be distributed during MPOA discovery.
4. A new request/response/trigger procedure between MPCs and MPSs can be defined to distribute this information.

The first method greatly increases the number of policy clients, and may affect the scalability of the policy distribution protocol. Additionally, it requires MPCs to have an IP address to run the policy distribution protocol (which is an explicit non-requirement of MPOA v1.0). However, this method requires no extensions in the behavior of any other LANE or MPOA component. In methods (2) through (4), new TLVs are defined to carry the mappings. The TLVs are carried in the appropriate LANE or MPOA control frames. The second method requires some out-of-band mechanism to distribute policy to the LANE Configuration Server, which then distributes it to the MPCs. With this method, there is a difficulty when an MPC has different policies when dealing with different MPSs. The third method is dynamic, but may be limited by the size of the LANE control frame (which is at most 1516 octets). The final method is the most generic, but requires a new type of MPOA control frame, a policy request/response frame. With the last method, the MPC must obtain the policy relevant to an IP destination before establishing a shortcut to that destination.

Methods (3) and (4) have the advantage that the distribution can be dynamic (i.e., the mappings can change and the changes are automatically propagated to the MPCs).

In choosing the distribution method, attention should be placed on the consideration for edge policy. The distribution of edge policy, in the form of filter specifications, differs from the distribution of the mappings between DSCPs and ATM QoS parameters in several aspects. First, the sheer amount of information in the collection of filter specifications is much greater. There are a limited number of DS Codepoints, but the number of possible filter specifications is much greater. Second, it is unlikely that a single MPC will need all of the filter specifications at any time. An MPC needs only those filter specifications that apply to traffic being forwarded by that MPC. Third, filter specifications are generally more dynamic and changes in filter specifications must be propagated in a timely fashion.

Proposed Solution for Policy Distribution

First of all, we assume that filter specifications can apply to a range of addresses (as specified by either a upper and lower bound, or a destination and mask). Filter specifications that apply to a single destination are better distributed to MPCs when traffic to the destination is detected. Filter specifications that apply to all destinations may be better distributed to all MPCs immediately upon device detection (without waiting for traffic to be detected). Between the two extremes is a range of possibilities.

As such, it is best left to the MPS to determine when filter specifications are distributed to MPCs. MPSs must ensure that all filter specifications that apply to a particular destination are distributed to an MPC before it initiates a

shortcut to the MPC. An MPS may choose any policy distribution algorithm it desires as long as this requirement is met. Thus, an MPS may choose to distribute all filter specification to an MPC upon device detection. Alternatively, an MPS may distribute filter specification to an MPC only upon flow detection. Many intermediate possibilities exist.

Diff-Serv PHB Policy Distribution and Specification

All MPSs (both Diff-Serv interior and edge nodes) need to forward to their MPCs the association between each DSCP and a set of ATM QoS and traffic parameters (an ATM traffic profile), which will be used for signaling shortcut VCCs. They must also provide any necessary mapping to ATM header marking, such as how to map the drop preference bits used in Assured Forwarding for the three drop preference levels to CLP bit for the ATM cells.

We propose that

- A new request/response/trigger procedure between MPCs and MPSs for distribution of Diff-Serv PHB policy. An initial request/response should be made immediately after an MPS and an MPC discover each other.

The PHB policy specification should contain the following fields:

Table 1. Contents of PHB policy specification

Name	Description
DSCP mask	Identify which of the 6-bit DS field are significant
DSCP	Bit pattern for Diff-Serv Codepoint

CLP marking	How a packet's drop preference indicated by this DS Codepoint should be mapped to the CLP marking in ATM header.
ATM Signaling IEs (one or more IE, each with three sub-fields)	As defined in UNI specifications
IE Identifier	
Length	
Contents of IE	

Each PHB policy specification TLV thus identifies one PHB through the pair of DSCP and DSCP mask. The ATM signaling IE fields are used when a new VCC needs to be established for carrying flows with this DSCP. The CLP marking field determines how the packet-to-cell adaptation should mark the ATM cells. As an example, a PHB group called Assured Forwarding (AF) utilizes 12 Codepoints for 4 AF classes, each with 3 drop precedences. Suppose the 4 AF classes are to be mapped to 4 virtual classes within the ATM UBR category, which are denoted here as Profile 1 through Profile 4. Also suppose the medium and high drop precedences are mapped to CLP=1. The PHB policy specification for each of the 12 DSCPs may be:

Table 2. An example of how the PHB policy specifications express the PHB-to-ATM QoS mapping.

DSCP Mask	DSCP	ATM Signaling IEs	CLP marking
111110	010000 (class 1, low)	Profile 1	0
111110	010010 (class1,	Profile 1	1

DSCP Mask	DSCP	ATM Signaling IEs	CLP marking
	medium)		
111110	010100 (class 1, high)	Profile 1	1
111110	011000 (class 2, low)	Profile 2	0
111110	011010 (class 2, medium)	Profile 2	1
111110	011100 (class 2, high)	Profile 2	1
111110	100000 (class 3, low)	Profile 3	0
111110	100010 (class 3, medium)	Profile 3	1
111110	100100 (class 3, high)	Profile 3	1
111110	101000 (class 4, low)	Profile 4	0
111110	101010 (class 4, medium)	Profile 4	1
111110	101100 (class 4, high)	Profile 4	1

Diff-Serv Edge Policy Distribution and Specification

An MPS at a Diff-Serv edge node must also distribute its filter specifications to its MPCs. As mentioned before, the solution must account for (1) the total volume of the filters, (2) their dynamic nature with policy changes, and (3) the fact that not all MPCs need all filters. An MPC only need those filters relevant to the packets it is forwarding.

Hence a solution capable of dynamic filter specification downloading is most desirable. We propose

- A new request/response/trigger procedure between MPCs and MPSs for distribution of packet classification and treatment policy.

In the following we outline one approach for such procedure, with the objective of ensuring that all filters relevant to an IP destination are available at an I-MPC prior to any shortcut establishment for that destination by the I-MPC. This assumes that the I-MPC is associated with a Diff-Serv edge router:

1. Initial filter specifications: An MPC requests from the MPS an initial filter set upon MPOA discovery. The MPS may return some or all filters specifications at its discretion.

2. Other filter specifications: An MPC initiates filter request/response for a specific IP destination address after the number of packets it receives for this destination exceeds a threshold and before it initiates an MPOA Resolution Request for this destination. The MPC may then use a second threshold to determine when to initiate a MPOA Resolution Request.

3. Policy changes: The MPS triggers an MPC to issue a filter request when there is a policy change. An MPS only needs to update those MPCs to which the filter specification applies.

We propose that

- The filter specifications from MPS to MPC used to communicate Diff-Serv edge policy should contain the following components and sub-fields:

Table 3. Contents of edge policy specification

Name	Description
Filter Operation	Specify what operation to be applied with this filter.
Filter Preference	Priority the filter relative to other filters in the case of multiple filter matches.
Filter Criteria (contains the following sub-fields)	Criteria used for packet classification
Ingress Interface	
Source IP Address	
Destination IP Address	
Carried Protocol	
Source Port	
Destination Port	
DS Field	
Filter Actions (one or more variable-length components expressed in TLV)	A flag identifying the action (such as shaping, policing, etc.), possibly followed by a number of variable-length parameters associated with the action (e.g., policing parameters).

Filter Operation indicates to the MPC what to do with this given filter. For example, operations may include install, delete, update, replace, enable, or disable. Edge policy may be specified such that a packet satisfies multiple filter

criteria cached at an MPC. To identify the appropriate filter actions to use, the filter specifications should be communicated with a relative preference.

Configuration and Discovery

To provide for interoperability with MPOA v1.0 implementation, we propose conveying Diff-Serv and Int-Serv capabilities in the discovery procedure. An absence of associated TLV(s) indicates that the device does not support the specific QoS capability.

Proposal:

- New TLV(s) should be added to the LE_ARP messages to indicate the QoS capabilities of the device during discovery.

Furthermore, other TLV(s) may also be added to indicate specific supports. For example, DS parameter distribution (e.g., service weights in a CBQ), methods of filter specification distribution, whether policy distribution is enabled, whether an MPS is a DS edge router, etc.

Flow Detection

As mentioned in Issue 2 above, the definition of "shortcut eligible flow" needs to be refined with QoS. We propose to adopt the following:

➢ For Diff-Serv traffic, flow detection uses a <MPS ATM Address, IP Destination Address, DS Codepoint> tuple. Flow detection definition and algorithms must be adjusted for the Diff-Serv codepoint.

In other words, ingress cache entries in MPCs should be extend d to monitor the DS field. With this definition, packets with the same IP destination but different DS Codepoints will be treated as separate flows as they correspond to separate ingress cache entries. Note that this is applicable whether the I-MPS is an edge or an interior node in its Diff-Serv domain. If multi-field packet classification is required, it must be done prior to the ingress cache lookup.

With dynamic edge policy distribution, the policy associated with a destination may not be at an ingress MPC prior to the initial arrivals of packets for that destination. This can be addressed by a two-stage flow detection. In the first stage, the I-MPC counts the number of packets with a particular destination address and MPS as if they are best-effort packets (DSCP = default). When this count exceeds a threshold, the MPC initiates a filter request. Upon receiving the applicable filter specifications, the I-MPC can perform flow detection including the DS Codepoint, based on the filter specifications returned.

Assume that an MPC is performing the edge device functions for a particular MPS and that it has requested and received the filter specifications applicable to a certain packet. When another matching packet arrives at the I-MPC, it is matched against the filter specifications to determine its DS Codepoint. The result, together with the destination address and MPS, is used to match with the ingress cache entries, or to create a new entry if one does not already exist. If a match is found and a shortcut VCC of the appropriate QoS is available, the I-MPC applies the corresponding filter actions before forwarding the packet over the shortcut. Otherwise, a flow detection counter is incremented and the packet is forwarded to the I-MPS, and no edge functions are performed by the MPC. If the configured flow threshold is exceeded, a shortcut with the appropriate QoS class is initiated. If the MPC is not performing edge router functions, the DS Codepoint of the packet is used to match the ingress cache entry.

Target Resolution

To address the needs for conveying DS Codepoints of flows to the E-MPCs (i.e., Diff-Serv Issue 3), we propose adding new extensions to the MPOA control message formats. Specifically, we propose that:

- A new MPOA QoS Extension is defined for the MPOA Resolution Request/Response to carry flow identification information. A possible format is as follows:

Field	Usage
Type	This field must be set to 0x---- (TBD) for MPOA DSCP Extension
Length	variable
Value	Filter criteria identifying traffic that will use this VCC

The propagation of this extension in other message formats (e.g., from NHRP Resolution Reply to MPOA Resolution Reply) is the same as other extensions. Note that this extension is not intended to be adjusted by intermediate routers but is intended to provide information to the egress MPOA devices so that the correct DLL and tag can be chosen for egress traffic.

In this case one DSCP extension is associated with each instance of target resolution (i.e., each unique <MPS ATM Address, IP Destination Address, DS Codepoint can initiate one target resolution). Hence if an E-MPS/E-MPC is capable of choosing a unique egress cache tag and/or a unique DLL header for each DS Codepoint, the values in the Egress Cache Tag Extension and/or the

DLL Header Extension in its MPOA Cache Imposition Reply will be uniquely associated with the Codepoint carried in the DSCP Extension.

Signaling and Connection Management

For Diff-Serv traffic, the DS Codepoint serves as the label that associates a Diff-Serv flow with a QoS VCC shortcut. In other words, an I-MPC should only forward a flow with a given DSCP over a VCC shortcut if the VCC is established using the signaling IEs given in the PHB policy specification for this DSCP.

CLAIMS

What is claimed is:

1. A method of providing Internet Protocol differentiated services using Multiprotocol Over ATM (MPOA) shortcuts providing for policy distribution through MPOA.

2. The method as recited by claim 1 wherein MPOA clients execute a policy distribution protocol to receive policy.

3. The method as recited by claim 1 wherein policy is distributed to MPOA clients by a LAN emulation configuration server.

4. The method as recited by claim 1 wherein policy is distributed to MPOA clients during a discovery process.

5. The method as recited by claim 1 wherein the policy is distributed to MPOA clients by MPOA servers using a request response protocol.

6. A method of distributing policy parameters in a network comprising the steps of:

   a) distributing policy information to MPOA servers; and

   b) distributing the policy information to MPOA clients.

7. The method as recited by claim 6 wherein the policy is used by MPOA clients to determine mapping of packets to differentiated service code points, to determine resource allocation and to determine packet metering, policing, marking and shaping.

8. A method of determining whether an IP flow is eligible for an MPOA shortcut comprising:

   a) a device determining the number of packets N sent with a DS code point P during a period of time T; and

   b) if N exceeds a value, the device determining the IP flow is eligible for an MPOA shortcut.

9. The method as recited by claim 8 wherein the device stores DS code point information in an ingress cache.

10. A method of determining whether an IP flow is eligible for an MPOA shortcut comprising:

    a) tracking packets sent with a DS code point by a device; and

    b) determining eligibility for an MPOA shortcut based on a predetermined criteria associated with tracking of packets sent with the DS code point.

11. The method as recited by claim 10 wherein the tracking comprises determining the number of packets N sent with the DS code point during a period of time T.

12. The method as recited by claim 10 wherein the step of determining eligibility comprises determining if the number N exceeds a value.

13. The method as recited by claim 10 wherein ingress cache entries in the device include DS code point information.

14. A device for routing information in a network, the device including an ingress cache, the ingress cache storing DS code point information for packets received by the device.

15. A method of distributing filter specifications from a MPOA server to MPOA clients in a network comprising:

a) distributing filter information which applies to a single IP destination to an MPOA client when traffic to the IP destination is detected at the MPOA client;

b) distributing filter information which applies to a range of IP destinations to an MPOA client in accordance with rules for the MPOA server; and

c) distributing filter information which applies to all IP destinations when a new MPOA client is detected.

16. The method as recited by claim 15 wherein the filter information which applies to a range of IP destinations is distributed to a particular MPOA client before initiation of a shortcut for one of the IP destinations by the MPOA client.

17. The method as recited by claim 15 wherein the filter information which applies to a range of IP destinations is distributed when a new MPOA client is detected.

18. The method as recited by claim 15 wherein the filter information which applies to a range of IP destinations is distributed when traffic to one of the IP destinations is detected at an MPOA client.

19. A method of detecting whether a device is capable of performing differentiated services comprising:

a) receiving a message from a device;

b) determining if the message includes elements encoded in type length value format indicating differentiated service capability; and

c) if the message includes such type length values, determining the device is capable of performing differentiated services.

20. The method as recited by claim 19 wherein the type length values indicate whether the device is an IP Differentiated Services edge device.

21. The method as recited by claim 19 wherein the type length values indicate whether the device has policy distribution capabilities enabled.

22. The method as recited claim 19 wherein the type length values Indicate whether the device has filter specification distribution enabled.

23. The method as recited by claim 19 wherein the type length values indicate whether the device has differentiated service parameter distribution enabled.

24. A method of updating filter specifications from an MPOA server to MPOA clients in a network comprising:

a) the MPOA server receiving filter specification updates from a policy server;

b) the MPOA server triggering MPOA clients to initiate filter request;

c) the MPOA clients initiating filter requests upon receiving a trigger from the MPOA server; and

d) the MPOA clients updating its filter specifications upon receiving filter responses from the MPOA server.

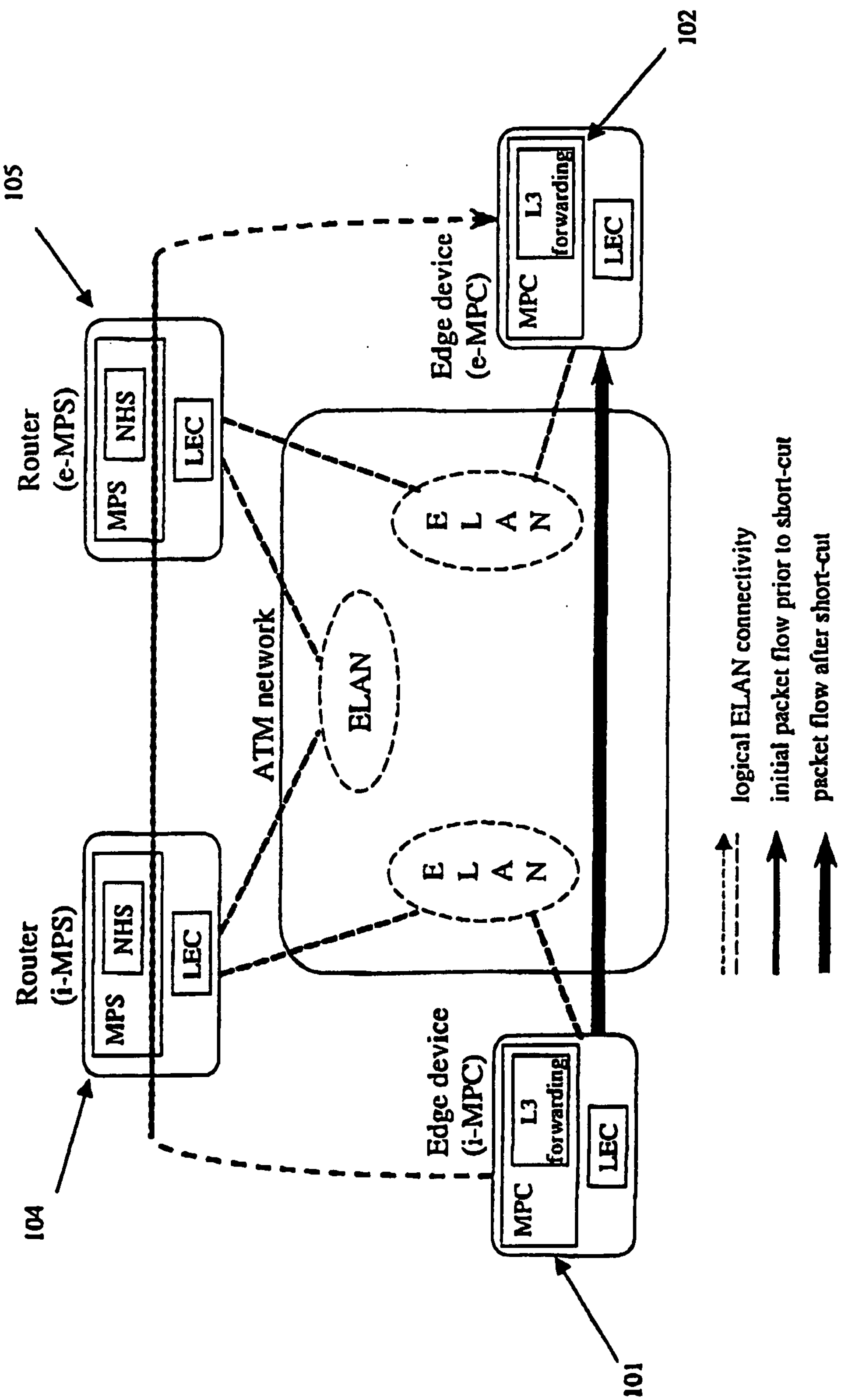

105
102
Router (e-MPS)
MPS
NHS
LEC
Edge device (e-MPC)
MPC
L3 forwarding
LEC
ATM network
ELAN
E L A N
E L A N
Router (i-MPS)
MPS
NHS
LEC
Edge device (i-MPC)
MPC
L3 forwarding
LEC
104
101
logical ELAN connectivity
initial packet flow prior to short-cut
packet flow after short-cut

FIGURE 1

INTERNATIONAL SEARCH REPORT

Interna. al Application No
PCT/US 99/26902

A. CLASSIFICATION OF SUBJECT MATTER
IPC 7 H04Q11/04

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 7 H04L

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	GILLHUBER A: "KURZE WEGE DURCHS NETZ. MULTI-PROTOCOL OVER ATM UND IP-SWITCHING" CT MAGAZIN FUER COMPUTER TECHNIK,DE,VERLAG HEINZ HEISE GMBH., HANNOVER, no. 11, 1 November 1997 (1997-11-01), pages 334-336,338,34, XP000704832 ISSN: 0724-8679 page 336, column 4, line 18 - line 48 page 341, column 2, line 40 -page 342, column 4, line 11 * lower figure * page 346	1-7
X	EP 0 866 630 A (NIPPON ELECTRIC CO) 23 September 1998 (1998-09-23) column 9, line 37 -column 14, line 11; figure 13B	8-18,24

-/--

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

° Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
20 April 2000	03/05/2000

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL – 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Gregori, S

1

Interns al Application No
PCT/US 99/26902

C.(Continuation) DOCUMENTS CONSIDERED T BE RELEVANT

Category *	Citation of document, with indication,where appropriate, of the relevant passages	Relevant to claim No.
X	CHEN X ET AL: "EVOLUTION OF ATM INTERNETWORKING" BELL LABS TECHNICAL JOURNAL,US,BELL LABORATORIES, vol. 2, no. 2, 21 March 1997 (1997-03-21), pages 82-110, XP000695170 ISSN: 1089-7089	19,20
A	page 101, column 1, line 19 -page 102, column 2, line 25	21-23
A	WO 98 24208 A (JONES RICHARD ;ZANCANI LEO (GB); DONKIN RICHARD (GB); LAURIE BEN () 4 June 1998 (1998-06-04) page 47, line 7 - line 25	1-7, 15-18,24

1

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No
PCT/US 99/26902

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
EP 0866630 A	23-09-1998	JP 2947201 B	13-09-1999
		JP 10229401 A	25-08-1998
		JP 11032047 A	02-02-1999
		CA 2229652 A	14-08-1998
WO 9824208 A	04-06-1998	AU 5062698 A	22-06-1998
		EP 0940024 A	08-09-1999
		GB 2319710 A,B	27-05-1998